

Introduction

JOSÉ MORAIS AND SÃO LUÍS CASTRO

Things, qualities, actions, relations, transformations, spatial and temporal references, all that, even mind properties, are matter for thought through words. Words are our universal way of putting sense in our relation with the world, with the others and with us, our most powerful control arm. Words have meaning, but also form and beauty. Words are particles, both corpuscles and waves, of our mental substance. The words in our head are what psychologists have called the mental lexicon – the mental representations of words –, and the study of the mental lexicon and of lexical processing is therefore a central topic in psychological science.

Words are not simple to study. First, they may have a complex morphological organization. Second, many languages contain several hundreds of thousands words, and each adult native speaker knows at least 20,000 base words (Gaskell & Ellis, 2009). Third, the mapping of words to their referents includes many cases of multiplicity (several words for the same referent, one word for several referents) and their frequencies vary quite largely. Fourth, more fundamentally and certainly not finally, words are complex configurations of articulatory gestures producing complex acoustic-phonetic patterns, analyzable into complex phonological structures, and recoded into complex graphemic and graphic forms, and all this is represented in our mind as units of information and in our brain as neural events.

Learning words is a hard task, especially in the first years of life, when the very notion that words refer to something must still be grasped and when the phonology of the mother tongue, including its rhythm and melody (the prosody, more precisely the changes in fundamental frequency, intensity and duration of speech), is just becoming familiar and has to be acquired. The continuous acoustic stream must be segmented to signal possible word boundaries, despite changes in speaker, affect and context, tremendous variations in acoustic and phonetic characteristics, and independent instances are to be categorized as corresponding to an abstract word. Prosodic boundaries, statistical information corresponding to frequent and phonotactic sequences, and pitch accent may be assumed to play a crucial role in structuring the linguistic input into units corresponding to words or from

which words are eventually derived. Learning words also requires some exemplars to be perceived as novel units, distinct from those already known, thus counteracting the assimilation tendency inherent to recognition within a lexical basis in which neighborhood density is constantly increasing and allowing connections with pre-existing words to be created. Learning the written form of words raises different problems. In the particular case of the alphabetic writing system, learners must understand the principle of grapheme-phoneme correspondence beneath it, master the rules of the orthographic code of the language, and store in a specific memory – the orthographic lexicon – the formal but nevertheless abstract representations of written words. All these steps are more hardly reachable when the code rules are complex, and the unpredictable exceptions to these are numerous.

The present special issue of the *Journal of Portuguese Linguistics* intended to address lexical processing. This is a rather broad topic and we were uncertain about the distribution of articles according to sub-topics. When the reviewing process was completed, we found that among the eight manuscripts accepted, seven (six of them had babies or children as participants) concerned some acquisition issue and four deal with prosody. Most likely indeed, our way of approaching potential authors presented a strong bias. Notwithstanding, the interest for lexical acquisition and prosodic processing seems to increase in the last years in the psycholinguistic research field. We had a look at ISI for papers published since 1987. We found that among the 4270 entries responding to the keyword “lexical processing”, those published at the beginning of the century (in 2000 and 2001) corresponded to 9.6% and those published in 2009 and in the first half of 2010 were 15.2%, a relative increase of 58% that testifies of the development of our discipline. Remarkably, in the same periods the percentages of scientific articles published were, respectively, 7.6% and 17.3% for those responding to the keyword “lexical acquisition” and 6.3% and 17.1% for those related to “prosodic processing” (relative increases of 120% and 170%). This suggests that the weight of research on lexical acquisition and on the role of prosody in lexical processing in the present special issue is not entirely at odds with the current trends in language studies.

The four initial papers describe studies on prosody, three involving children (Salselas & Herrera; Mersad, Goyet, & Nazzi; Millotte, Morgan, Margules, Bernal, Dutat, & Christophe), and the fourth on how it can be used by adults, together with statistic cues, to segment speech (Fernandes, Ventura, & Kolinsky). Then, two papers are concerned with the acquisition of morphology, one examining pre-readers (Corrêa, Augusto, & Castro), the other comparing good to poor readers (Duncan, Gray, Quémart, & Casalis). The two final papers deal with the processing routes that eventually activate lexical representations, one testing adults with speech (Serniclaes, Beeckmans, & Radeau), the other examining the course of reading and spelling development (Serrano, Genard, Sucena, Defior, Alegria, Mousty, Leybaert, Castro, & Seymour). Several languages are contemplated in these papers: Portuguese (in two of three, both

the Brazilian and the European dialects), English, French, an artificial language, which included Portuguese-like stress patterns, and in two papers either English and French, or French, English and Spanish are compared.

Salselas & Herrera examined the vocal (speech and singing) material directed to infants, from which these are supposed to begin extracting lexical information. The lexicons of European and Brazilian Portuguese are roughly the same, however, there are important prosodic differences between the two variants. Recordings of caregivers interacting with their babies up to 18 months in several speech (affection, disapproval, questioning) and singing pragmatic classes, were segmented into durational units, analyzed into pitch- and rhythm-related descriptors, and finally submitted to a discrimination model in three machine learning experiments. This model proved to be able to discriminate the two variants in speech and, to an apparently greater extent, in singing, mainly on the basis of rate and both vocalic and consonantal (but not syllabic) durations. The two Portuguese variants have indeed distinct rhythm patterning, which appears to be present even in interactions with babies. While the variants discrimination relied only on rhythm-related descriptors in the case of speech, it involved also pitch-related descriptors in the case of singing. However, as expectable, pitch- but not rhythm-related descriptors, were efficient in discriminating between speech pragmatic classes within the same variant. Most important, this study may provide an inspiring piece of evidence on the hot issue of whether language and music share, or to what extent do they share, the same kind of computational resources during the preverbal developmental period.

How infants segment word forms from speech is a crucial – logically the first – issue regarding lexical acquisition. Mersad, Goyet, & Nazzi propose the “early rhythmic segmentation” hypothesis, based on the idea that infants begin by extracting the rhythmic unit of their native language, independently of any lexical information. This is possible given the young infants’ sensitivity to rhythm as a case of the well-documented early sensitivity to prosodic information. That hypothesis is integrative, valid for all languages. However, given the existence of rhythmic classes of languages, the segmentation procedures developed by the infants should be different for, in particular, stress- and syllable-based languages, well represented by English and French, respectively. The authors review empirical evidence that supports all these assumptions, mostly from American English-learning infants and both Canadian and Parisian infants native of French. Interestingly, an age delay of about 4 months is observed in the access to the corresponding rhythmic unit for English and both French dialects. That delay presumably arises from the specific difficulty created by the syllable unit for French-learning infants in operating whole word segmentation when words are polysyllabic. The French cross-dialectal studies suggest that in this process the distributional analysis of syllable order may provide a disambiguating segmentation cue. At the end of their paper, the authors remind us that European Portuguese, with both

syllable- and stress-timing properties, should be a relevant research case for the necessary cross-linguistic approach to the roots of lexical acquisition.

Millotte, Morgan, Margules, Bernal, Dutat, & Christophe also deal with the issue of how infants do segment the speech stream into word-sized units, but they focus on the role of the phonological phrase within a sentence in constraining lexical access. This phrase, which contains one or two content words and associated function words, is characterized by preboundary lengthening and presents peculiar prosodic variation. In adults, phonological phrases are processed as chunks facilitating or interfering with lexical access when, respectively, the lexical candidate is fully included in them or straddles a boundary. The present experiment used the head-turn preference technique to show that 16-month-old infants native of French behave like the adults: reinforced to respond to a bisyllabic target like “balcon”, they oriented more often to it when it was fully included in the prosodic unit than when it straddled the prosodic boundary. Conversely, reinforced to respond to “bal” they oriented less often to it when “bal” was the first syllable of “balcon” than when it was followed by a word beginning with “con-“. In other words, lexical access is guided by phonological phrase processing since this early age. Quite interestingly, however, no such effects appeared for 10-month-olds. For French, compared to English, there seems to be a delay of 3 to 6 months in the development of the role of phonological phrases in lexical access, perhaps due to a greater difficulty in accessing whole words in a syllable- than in a stress-based language, or, as the authors also acknowledge, to different styles of motherese in the two linguistic cultures. Note that the two reasons are not exclusive. The authors compared the vocabulary development in the USA and in France and found again a delay, though shorter, for the French children.

Unfamiliar, artificial languages provide an appropriate material for comparing the contribution, presumably interactive, of different types of speech segmentation cues to what could be “words” in those languages. This is the strategy employed by Fernandes, Ventura, & Kolinsky in a study with adult listeners, in the occurrence Portuguese ones. Their basic hypothesis was that the weighting of segmentation cues depends on domain-generalities, i.e. being used not only for speech but also for non-linguistic auditory material and even for visual sequences, and/or on universality, i.e. being language-independent or -dependent. Transitional probabilities are a universal cue, and their use seems to have a phylogenetic origin. Intonational or prosodic phrases with right edges, due to the slowing down of the articulators within a breath group and characterized by final lengthening and low pitch, constitute a language-independent cue; and lexical stress constitutes a language-dependent cue. The resilience of these cues to physical degradation through white noise superimposition, and their contribution to a redundancy gain when congruent with each other, provided two distinct ways of verifying their role. The task consisted, after a familiarization phase with a stream of the artificial language,

in choosing the “word” between two stimuli. Both transitional probabilities and intonational phrases were highly resilient to physical degradation. Duration, which was used as a correlate of lexical stress, had no impact with intact speech and its role was highly dependent on listening conditions. The results, only partially consistent with the basic hypothesis, provide nevertheless relevant data for future research.

Whereas the processing of prosody and of phonological phrases is unanimously recognized as developing quite early, the time of emergence of morphological knowledge, in particular as regards the production of gender agreement in determiner phrases, seems to remain a controversial issue. It has been proposed that gender agreement is already acquired by three years of age as an important part of our language system; according to an opposite view, however, it would result from rather late strategic cue-based learning. Following a previous study, and examining 2- to 4-year old Brazilian and Portuguese children in an elicited production task that required them to ascribe gender to novel objects designated by pseudo masculine/feminine nouns, Corrêa, Augusto, & Castro present here evidence that clearly supports the former position. In this situation, the determiner (the masculine/feminine article) provided the only relevant information to produce the noun gender, and, as a matter of fact, most of the time the children responded in the way that was consistent with the determiner. The children thus seem to have acquired an algorithm procedure to generate the noun gender according to the gender of the determiner. This procedure allows the novel noun to be represented in the lexicon with the gender of the associated determiner, which becomes an intrinsic feature of it. Interestingly, the mismatch between the gender of the determinant and the noun end was more disturbing when the determiner phrase was feminine (for example “a bido”) than when it was masculine (“o daba”), consistent with the idea that the feminine gender is the marked form.

The morphological representation in the mental lexicon is also the main topic of the paper presented by Duncan, Gray, Quémart, & Casalis. The decompositional perspective is adopted, i.e. the recognition of words involves their decomposition into morphemes, represented in either a free or a bound manner, and their later combination or integration to assess the lexical representation of the whole word. The specific issue of this study is the extent to which morphological ability demonstrated in oral language contributes to decoding ability in English readers of grades 3 and 4, and whether poor readers manifest both poorer morphological awareness and poorer recognition of morphemes in reading. Morphological awareness was assessed by asking the children to complete a sentence with a word that was a derivation of a word included in the sentence. A task of lexical decision to written items, in which these were chosen in such a way that they included or not an embedded word and a suffix, provided an estimation of the use of morphemes in reading. The results were consistent with the decompositional theory. Poor readers displayed increased accuracy in accepting a target word with an embedded

word, even though, contrary to what good readers did, in this case they were not sensitive to the presence of an orthographic suffix. The detection of embedded words may be an important, but still elementary, part of the developmental process of word reading. Contrary to expectations, however, neither group showed a relation between morphological decomposition in reading and morphological awareness in the sentence completion task, which suggests that children at this age have resort to morphological skills that are still strongly task-dependent.

The influence of lexical information on a in-principle lower process, namely phonetic categorization of a spoken item, was revisited by Serniclaes, Beeckmans, & Radeau. It was indeed found at exactly thirty years ago that, in the choice between a word and a nonword differing in only one feature, like voiced-unvoiced, there is a lexical identification shift, i.e. perceptual responses indicate a displacement of the neutral point in the acoustic continuum towards the nonword. Both an interactive top-down model and a model assuming a lexical decision bias on two independent processes, one phonetic, the other lexical, can account for this shift. While being the first to look for this effect in French, the main reasoning of the authors was to assess its modulation by the length of the word in a more systematic way than it had been done in the past. Quite interestingly, a U-shaped relationship was found, with a clear effect with monosyllables (for example, *dame-tame*), similar to the one usually obtained in English, no effect with bisyllables (*madame-matame*), contrary to the expectations in face of the literature, and an effect in polysyllables (*hirondelle-hironzelle*) that was similar in size to the one with monosyllables and stronger than the difference obtained with bisyllables. In addition to showing that the lexical identification shift can be obtained in quite different languages, the authors were able to argue that the length effect, correlated with other lexical variables all concurring for lexical evidence to increase with word length, is not a simple matter of lexical evidence. A second mechanism would operate in the time limits of echoic memory, which the authors call “contrastive scanning” of the phonetic input, and which would be responsible for the drop of the effect with bisyllables. This is a provocative proposition that will certainly inspire further work.

In the last contribution to the present special issue on different aspects of lexical processing and its acquisition, Serrano, Genard, Sucena, Defior, Alegria, Mousty, Leybaert, Castro, & Seymour examined reading and spelling development longitudinally, throughout the first grade (October, February, and May), in three Romance languages, French (Belgium), Portuguese, and Spanish. The main variables that were assumed to contribute to an expected cross-linguistic effect were the degree of bidirectional consistency between orthography and phonology (the greatest in Spanish and the smallest in French), including an asymmetry between the complexity of grapho-phonological and phono-graphic conversion rules, which is the greatest in French and the smallest in Spanish). In their analyses of the data, the authors

contemplated also the role of vocalic reduction (the greatest in Portuguese) and of the number of vowels (much smaller in Spanish than in both French and Portuguese). The instructional method was similar in all languages: teaching of letter-sound and grapheme-phoneme correspondences in a semantic context. Whereas letter knowledge at time 1 did not differentiate Spanish from Portuguese children, and only barely from French children at time 2, the Spanish learned to read and spell words at a much faster pace than children native of the other two languages. For reading, the Spanish superiority, already strong in accuracy, was particularly dramatic in speed. There was evidence that all the factors mentioned above contribute to these cross-linguistic differences. Importantly, across the three languages, early letter-sound knowledge was found to be a good predictor of both lexical (word reading) and sublexical (pseudoword reading) processes.

We were happy to collect this set of articles to the *Journal of Portuguese Linguistics*. Though not focused on a single problem, they present, by groups, remarkable convergence regarding the contribution of prosody, phonology and morphology to lexical processing, with an emphasis on lexical acquisition and the spoken language but without entirely neglecting adult processing and the written language. Cross-linguistic and even cross-dialectal inquiry was a further characteristic of most of the studies, and some of these involved Portuguese. We thank the authors, obviously, as well as the experts. These, remaining anonymous, read and commented the manuscripts very carefully and contributed significantly to improve the quality of many of them.

Reference

- Gaskell, M.G., & Ellis, A.W. (2009). Word learning and lexical development across the life span. *Philosophical Transaction of the Royal Society B*, 364, 3607-3615.

José Morais
Unité de Recherche en Neurosciences
Cognitives (UNESCOG)
Université Libre de Bruxelles (ULB)
50, Avenue F.D. Roosevelt
B-1050 Bruxelles, Belgium
jmorais@ulb.ac.be

São Luís Castro
Faculdade de Psicologia e Ciências
da Educação
Universidade do Porto
Rua Dr. Manuel Pereira Silva
P-4200-392 Porto, Portugal
slcastro@fpce.up.pt